of aggression, observational studies are limited in what they can address. Interactive animal studies on the neurocircuitry of aggressive behavior have generated a wealth of data which are translationally relevant for disorders involving altered aggressive responses, such as post-traumatic stress disorder (PTSD) (Ferris et al. 2008). The comments by Schilbach et al. on the importance of reward circuits and social interaction suggest that second-person neuroscience is a valuable addition to the study of depression as well. Anhedonia and attenuated social interaction are common features of depression, and the second-person neuroscience method can address both features in studies of depression in males and females. In humans, Lane Strathearn and others have shown that it is possible to record differences in fMRI responses to stimuli such as infant cry or infant pictures in healthy mothers (Strathearn et al. 2009), but the most relevant BOLD differences are from studies comparing healthy and depressed mothers (Laurent & Ablow 2012). Strong collaborations between animal and human researchers focused on developing ecologically and ethologically relevant second-person neuroscience experimental paradigms will advance the development of social neuroscience.

Social perception and "spectator theories" of other minds

doi:10.1017/S0140525X12002014

Søren Overgaard[a] and Joel Krueger[b]

[a]*Center for Subjectivity Research, University of Copenhagen, 2300 Copenhagen S, Denmark;* [b]*Department of Sociology, Philosophy and Anthropology, University of Exeter, Exeter EX4 4RJ, United Kingdom.*

s.overgaard@hum.ku.dk **http://cfs.ku.dk/staff/profil/?id=259148**

j.krueger@exeter.ac.uk **www.joelkrueger.com**

Abstract: We resist Schilbach et al.'s characterization of the "social perception" approach to social cognition as a "spectator theory" of other minds. We show how the social perception view acknowledges the crucial role interaction plays in enabling social understanding. We also highlight a dilemma Schilbach et al. face in attempting to distinguish their second-person approach from the social perception view.

Schilbach et al. argue that their "second-person" approach to social cognition – which emphasizes the centrality of interaction and emotional engagement in facilitating social understanding – differs fundamentally from what they term "spectator theories." Under the latter heading they include not just "Theory Theory" (TT) and "Simulation Theory" (ST), but also what they aptly term the "social perception" approach. Although we suspect similar questions could be raised with respect to TT and ST, we focus on the question of whether it is true that the social perception approach is "committed to an observer epistemology" (sect 1.2, para. 2) and thus is a "spectator theory."

The social perception approach comes in different forms (see, e.g., Duddington 1918; Gallagher 2008; Hobson 2008; Scheler 1954; Stout 2012; Wittgenstein 1980). Common to all is the idea that we can sometimes perceive (and thus need not infer) the mental states of others (Krueger 2012; Krueger & Overgaard 2012; Overgaard 2012). But *contra* Schilbach et al., it is not clear that this commits the social perception approach to a "spectator theory of how we relate to other people" (sect 1.2, para. 2). For, first of all, there is no reason to construe the social perception theory as a general theory of "how we relate to other people." Not only is the theory fully consistent with the idea that social life is more a matter of interacting with, than merely observing others. Additionally, it can also maintain that when we perceive others' emotions and goals, say, we typically do so in order to facilitate or enable interacting with them. For example, I may perceive anger in another person's facial expressions and gestures; the former expresses specific features of the anger whereas the latter articulates its intensity and level of arousal (Ekman 1965). But this need not involve a static observational stance. Rather, I see this anger as affording (or perhaps constraining) further interactive possibilities. My social perception is thus typically part of an ongoing process in which I use what I see in others to further negotiate my engagement with them.

More importantly, the social perception account can maintain that the ability to perceive others' mental states may be a capacity that is developed and/or refined within interaction, for example, the face-to-face interaction characterizing early infant-caregiver exchanges (see, e.g., Hobson 2002; Legerstee 2005). Moreover, even in adult perceivers, our exercise of the ability may be greatly facilitated by current interaction with others – and may in some cases be hindered in the absence of interaction (see, e.g., Gallagher 2008). For example, certain socially-salient details – gestures, intonation, posture, shifts of spatial proximity, and so forth – may only become available for perception as an interaction unfolds over time. And when we perceive facial gestures (e.g., a smile or frown), we automatically attune to these gestures by involuntarily mimicking them (Dimberg & Thunberg 1998). As one of the target article authors notes in another paper, this mimetic response seems to indicate that "the process of perceiving faces always includes an 'enactive' element through which we engage with and respond to stimuli instead of a mere 'passive' perception of face-based cues" (Schilbach et al. 2008a, p. 45). The social perception approach is fully compatible with this observation.

But Schilbach et al. are right that at least standard versions of the social perception approach suggest that social cognition is "something that could in principle occur in the presence of a one-way mirror, where a detached observer reads out the mental states of another person" (sect. 2.1, para. 2). However, it seems highly plausible that "detached" social perception of the "one-way mirror" sort could *in principle* occur; indeed, it might be something that happens fairly frequently – as, for example, when we perceive the emotions expressed by people on TV, while watching one's fellow diners in a restaurant, or when viewing others from a distance.

So Schilbach et al. are faced with the following dilemma. Either they maintain that social perception is in principle impossible in the absence of current social interaction, as, for example, in the case of the one-way mirror. This is a very strong claim: A single instance of social perception without interaction (e.g., successfully perceiving a diner's expression of happiness *as such* across a crowded restaurant) would suffice to falsify it. Or else they acknowledge that the latter is a possibility. But then they are wrong to suggest that their "second-person" approach is distinguished from the social perception approach by the "spectatorial" nature of the latter. For, the sort of purely "spectatorial" social cognition that the social perception approach as such makes room for (e.g., recognizing emotional expressions at a distance or in the absence of interaction) is one that Schilbach et al. themselves want to accommodate.

In our view, it is crucial to recognize the broad scope of our social competence, encompassing situations in which we are active participants, as well as situations in which we are passive bystanders. Despite their justified emphasis on the primacy of interaction, Schilbach et al. recognize this complexity and attempt to develop their view accordingly. However, in doing so, they fail to show that the social perception approach to social cognition is genuinely committed to the sort of "observer epistemology" they see their own view as opposing.

Although we have here focused on the social perception approach, our commentary also raises the more general question of whether the theoretical commitments of existing approaches really are so opposed to the "second-person" approach championed by Schilbach et al. Defenders of TT, for example, have long emphasized the importance of social interaction for the development of desire and belief understanding (e.g., Repacholi

& Gopnik 1997; Ruffman et al. 1998). "Development in social cognition depends on two-way traffic between self and other," as Meltzoff et al. (1999, p. 19) put it more than a decade ago. So perhaps it is not just the social perception approach that in reality is less "spectatorial" than Schilbach et al. make it appear to be.

Interaction versus observation: A finer look at this distinction and its importance to autism

doi:10.1017/S0140525X12002026

Elizabeth Redcay,[a] Katherine Rice,[a] and Rebecca Saxe[b]

[a]*Department of Psychology, University of Maryland, College Park, MD 20742;* [b]*Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139.*

redcay@umd.edu krice@umd.edu saxe@mit.edu
www.dscn.umd.edu www.saxelab.mit.edu

Abstract: Although a second-person neuroscience has high ecological validity, the extent to which a second- versus third-person neuroscience approach fundamentally alters neural patterns of activation requires more careful investigation. Nonetheless, we are hopeful that this new avenue will prove fruitful in significantly advancing our understanding of typical and atypical social cognition.

We enthusiastically applaud the call for a second-person neuroscience as described in the target article by Schilbach et al., and are excited for the new insights this avenue of research will bring. In this commentary, we expand on the target article in two ways. First, we suggest increased emphasis on characterizing the differences in neural processing during an interaction as compared to observation. Second, we elaborate on the potential importance of this research to our understanding of autism.

Schilbach et al. argue that engagement in interpersonal interaction fundamentally changes cognitive and neural processing as compared to such processing during observation alone. For example, interpersonal interaction may recruit additional neural regions or systems that are not present during third-person observation. The authors present compelling preliminary pieces of evidence to support their theory, as well as many suggestions for future directions. Here we discuss several notable differences between observing a person and engaging with a person, which can make isolating the interaction component difficult. First, engagement with another involves a contingency (or back and forth) between participants rather than passive perception and as such includes an element of action. Second, and related to the first, when a response is required (as is common in an interaction as compared to observation) attentional demands may be higher. Third, the stimuli used to elicit the feeling of being in an interaction have different low-level characteristics than those that signal no interaction. A final possibility is that there is something special about being engaged with another that goes beyond the simple differences described above.

Controlling for these differences is important, and although Schilbach et al. devote attention to this problem, they do not address whether stimulus characteristics secondary to the interaction could drive differences in neural processing. For example, in one study participants are presented with a face either facing towards or away from them that makes either communicative or arbitrary facial movements (Schilbach et al. 2006). Ventral medial prefrontal cortex (vMPFC) and amygdala regions are recruited to a greater extent for the communicative facial expressions directed towards the participant but these regions may be sensitive to direct gaze and facial movement independent of social engagement. While these expressions are typically encountered in the context of an interaction, they are also seen in movies, TV, and pictures when the viewer is (presumably) detached. Although interaction is often a more ecologically valid social situation, it is an open question of how, once other factors are controlled for, this interaction fundamentally changes the neural correlates of social processing.

We have begun to address this question (Redcay et al. 2010) by borrowing a method from developmental psychology (e.g., Kuhl et al. 2003; Murray & Trevarthen 1985) in which participants are engaged in a simple, highly scripted interaction that is either conducted via live video feed ("face-to-face") or via video recording. The recorded conditions included one in which the same video from the live interaction was repeated and one in which a video of the experimenter from a different interaction was played. Crucially participants were told to continue to play along in the recorded condition even though the experimenter would not be able to see or hear them. These controls allowed for an examination of brain regions that were recruited during an interaction that could not be accounted for by differences in stimulus properties. Comparison of live and recorded conditions revealed that largely the same set of brain regions were engaged in both conditions. For example, robust recruitment was seen in the posterior superior temporal sulcus (STS) during both live and recorded conditions, which is not surprising given the STS's role in human action perception (e.g., Pelphrey et al. 2004; Saxe et al. 2004). Interestingly, the live condition showed increased activation of the posterior STS, and this extended more posteriorly into the temporo-parietal junction (TPJ), a region associated with theory of mind processing. Thus, this study offers support for differences in the magnitude of activation in brain systems between live, contingent interaction, and non-contingent interaction when stimulus characteristics are constant and some support for fundamental differences in the brain systems recruited. Future studies which control for action, attention, and stimulus characteristics (in addition to those proposed by Schilbach et al.) will be critical to disentangle where the "bookends" (sect. 4) begin and end; in other words, what are the differential effects of second- versus third-person approaches to social cognition on neural patterns of activation?

Characterizing these "bookends" is especially important for understanding autism, a developmental disorder characterized by impairments in social interaction, particularly in the intentional coordination of attention with others, or joint attention (e.g., Charman 2003; Mundy & Newell 2007). However, offline laboratory-based tasks often fail to find deficits in joint attention behaviors (Nation & Penny 2008; Redcay et al. 2012). Similarly, tasks tapping into belief representations demonstrate fairly typical performance (e.g., Senju et al. 2009) and even typical neural patterns of activation (e.g., Dufour et al. 2012). One possibility is that these third-person studies may be failing to capture the challenges of a real-time social interaction for a person with autism. A recent study of ours (Redcay et al., in press) compared patterns of activation during a real-time joint attention game between high-functioning adults with autism and typical adults. Whereas typical adults demonstrated selective recruitment of the left posterior STS and dorsal medial prefrontal cortex (dMPFC) during joint, as compared to solo, attention, the participants with autism revealed a pattern of reduced selectivity due to both hypoactivity during the joint conditions and hyperactivity in the solo condition. These data suggest a failure to modulate these brain regions according to whether the task required a social interaction. Importantly, the differential effects of second- versus third-person interaction might vary between typical and atypical populations, or change throughout development. This presents a major challenge to our understanding of the neurobiology of social processing in autism, but we are optimistic that a continued second-person neuroscience approach will reveal the mechanisms underlying real-world social difficulties in autism.